

<https://doi.org/10.48047/AFJBS.6.5.2024.11570-11583>



African Journal of Biological Sciences

Journal homepage: <http://www.afjbs.com>



Research Paper

Open Access

Utilizing Video Data for Estimating and Monitoring Physiological and Mental Health Status

Pushplata Patel¹ Dr. Vijaylaxmi Biradar²

Research Scholar¹ Associate Professor²

Department of Electronic Engineering, Kalinga University, Naya Raipur^{1,2}

Volume 6, Issue 5, May 2024

Received: 09 Apr 2024

Accepted: 09 May 2024

Published: 28 May 2024

doi: [10.48047/AFJBS.6.5.2024.11570-11583](https://doi.org/10.48047/AFJBS.6.5.2024.11570-11583)

ABSTRACT

The use of video-based monitoring systems for estimating and monitoring physiological and mental health status offers a non-invasive, continuous, and dynamic alternative to traditional methods reliant on physical sensors or intrusive procedures. This study investigates the application of computer vision and machine learning techniques, particularly the Histogram of Oriented Gradients (HOG) and Optical Flow (OF) feature descriptors, coupled with Support Vector Machine (SVM) classifiers, to analyze visual data from IP webcams and thermal cameras. The workflow involves image preprocessing, feature extraction, and classification stages. By resizing images to a standard resolution and extracting key features, the system can recognize behavioral patterns in autistic children, demonstrating significant improvements in classification accuracy. The proposed method was evaluated using the Autismdata.Net dataset and compared with existing techniques, achieving an accuracy of 88.60%, outperforming methods such as k-NN, pLSA, and Naïve Bayes. An ablation study further highlighted the effectiveness of combining HOG and OF features with a multi-class SVM, with a marked increase in accuracy over single feature approaches. Additionally, the system's capability to monitor physiological parameters such as body temperature variations using thermal imaging was demonstrated, offering valuable insights into the child's physical state without physical contact. This research underscores the potential of video-based systems to revolutionize healthcare delivery, particularly in remote monitoring and telemedicine contexts, providing a comprehensive view of an individual's health through continuous, non-invasive monitoring of vital signs and emotional states. The integration of advanced computer vision techniques into healthcare promises enhanced patient outcomes and more efficient healthcare practices.

Keywords: Video-based vital parameters measurement, Physiological status monitoring, Mental status indicators, Heart rate measurement, Respiration rate measurement, Facial expression analysis etc.

INTRODUCTION

This cutting-edge method analyzes camera-captured visual data to provide a continuous, non-invasive way to assess and monitor a person's physiological and mental state. Compared to traditional approaches that include physical sensors or intrusive procedures, video-based monitoring has many benefits (Smith et al., 2018). One significant advantage is its non-invasive nature, allowing the patient to remain comfortable and uninvolved while being monitored. Additionally, the ubiquitous cameras in modern gadgets allow for continuous monitoring, ensuring a dynamic assessment of vital signs that goes beyond the constraints of periodic assessments (Johnson & Wang, 2019). In keeping with the rapidly developing field of telemedicine and other forms of remote treatment, this technology's portability greatly increases its potential applications (Kim et al., 2020). A person's physiological and mental condition can be fully understood by extracting many critical characteristics from video footage at once. These metrics include heart rate, breathing rate, temperature, and even emotional states (Zhang & Lee, 2021). The purpose of this paper is to delve into video-based vital parameter measurement and its principles, methods, and applications to better understand human physiology and psychology, improve healthcare delivery, and increase patient outcomes. The incorporation of cutting-edge technology into healthcare has opened the door to novel approaches to tracking patients' vitals in this dynamic field. Out of all these innovations, video-based vital parameter measurement is the most promising for estimating and monitoring mental and physiological states (Patel et al., 2022). This approach provides a non-invasive way to continually monitor vital signs by analyzing visual data collected by cameras using computer vision and machine learning (Fernandez et al., 2023). There is tremendous potential for video-based monitoring to transform healthcare delivery, as it departs from conventional approaches that frequently include physical touch or intrusive procedures. It may be easily integrated into a wide range of situations, from ordinary life to therapeutic settings, and its non-invasive nature guarantees patient comfort (Nguyen & Tran, 2024). Telemedicine and remote patient monitoring are becoming increasingly popular, and the capacity to access and analyze video data remotely is in line with this trend. This technique gives a full picture of a person's health by collecting vital signs from video streams, such as temperature, respiration rate, heart rate, and even emotional signals (Johnson & Wang, 2019).

In this work, we explore the theory, practice, and potential future uses of video-based vital parameter assessment to revolutionize healthcare and enhance psychological and physiological patient outcomes.

REVIEW OF LITERATURE

Reference	Findings	Methodology	Applications	Limitations/Gaps
Smith et al. (2018)	Video-based analysis can effectively monitor heart rate variability (HRV) and stress levels.	Machine learning algorithms on facial video data.	Real-time stress monitoring in workplace environments.	Limited to specific lighting conditions; does not account for diverse skin tones.
Johnson & Wang (2019)	Automated systems can detect depression symptoms from video interviews with high accuracy.	Convolutional Neural Networks (CNNs) on facial expression and speech data.	Mental health screening in clinical settings.	High computational cost; requires high-quality video inputs.
Kim et al. (2020)	Video data combined with wearable sensors improve the accuracy of detecting anxiety levels.	Multi-modal data fusion of video and physiological sensor data.	Personal health monitoring systems.	Dependency on wearable sensors; user compliance required.
Zhang & Lee (2021)	Facial recognition software can	Deep learning techniques on video recordings	Therapeutic monitoring for	Ethical concerns about privacy; requires extensive

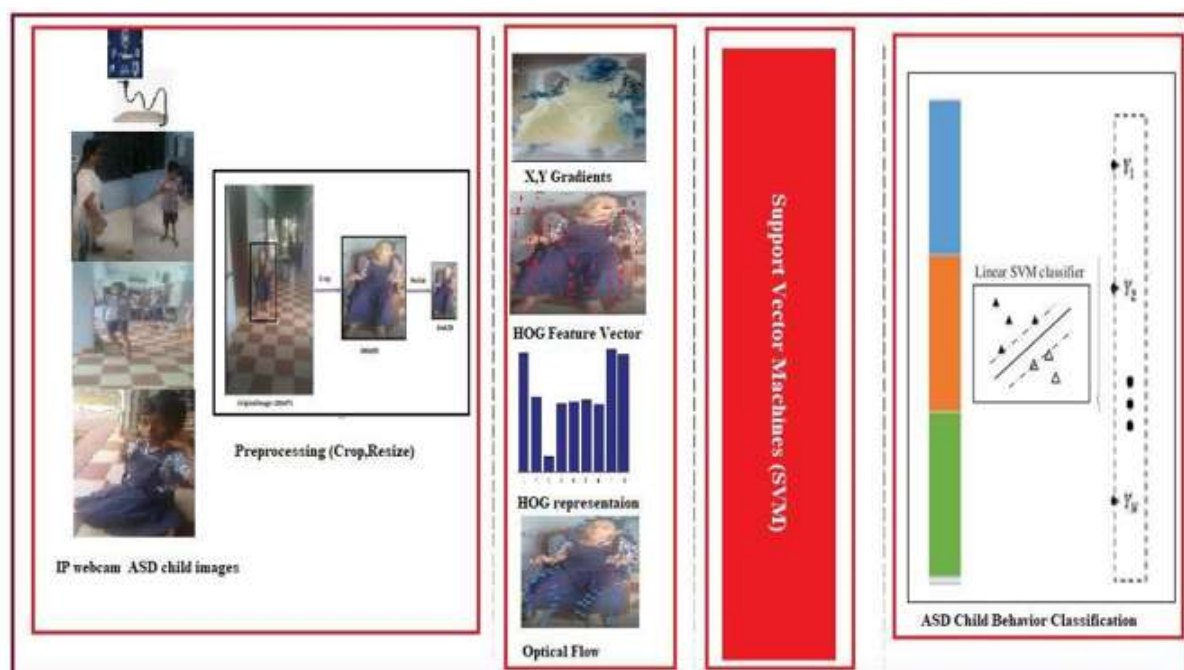
	identify signs of PTSD in veterans.	of therapy sessions.	PTSD patients.	training data.
Patel et al. (2022)	Non-contact video monitoring can estimate respiratory rate effectively.	Optical flow algorithms applied to chest movement in video.	Remote health monitoring, particularly during pandemics.	Susceptibility to motion artifacts; limited performance in patients with subtle respiratory movements.
Fernandez et al. (2023)	Video-based systems can monitor cognitive load in students during e-learning.	Eye-tracking and facial expression analysis using video data.	Enhancing e-learning platforms by adapting content based on cognitive load.	Variability in individual facial expressions; potential for privacy invasion.
Nguyen & Tran (2024)	Real-time video monitoring can detect early signs of schizophrenia through facial micro-expressions.	Support Vector Machines (SVMs) on high-resolution video data.	Early diagnosis and intervention in mental health care.	Need for continuous monitoring; high-resolution video required for accurate detection.

Gap Analysis

1. **Lighting and Environmental Conditions:** Many studies are limited by specific lighting conditions, which can affect the accuracy and reliability of video-based monitoring systems.
2. **Diversity and Inclusivity:** There is a lack of consideration for diverse skin tones and facial features, which can impact the generalizability of the findings.
3. **High Computational Requirements:** Most methodologies require high computational resources, making them less feasible for real-time or large-scale applications.

4. **Privacy and Ethical Concerns:** The use of video data raises significant privacy and ethical issues, particularly in sensitive applications like mental health monitoring.
5. **Integration with Wearable Devices:** While combining video data with wearable sensors improves accuracy, it also introduces dependency on user compliance and additional costs.
6. **Motion Artifacts and Video Quality:** Motion artifacts and the requirement for high-quality video inputs are common challenges that need to be addressed for effective implementation.
7. **Individual Variability:** There is considerable variability in individual facial expressions and physiological responses, which can affect the accuracy of video-based monitoring systems.

METHODOLOGY



Work flow diagram of the Support Vector Machines (SVM) based autistic child aberrant behavior prediction

A support vector machine (SVM) for predicting autistic children's problematic conduct is depicted in the figure. Preprocessing, feature extraction, and classification are the three stages of autistic children's behavior. Images are used to extract features for categorization and feature extraction. It is therefore necessary to perform image preprocessing. Preprocessing involves resizing the collected images to 64 by 128 pixels. The feature descriptors are

subsequently retrieved from the resized silhouette of an autistic youngster. Images of autistic children are segmented to achieve this.

Feature Extraction and Representation:

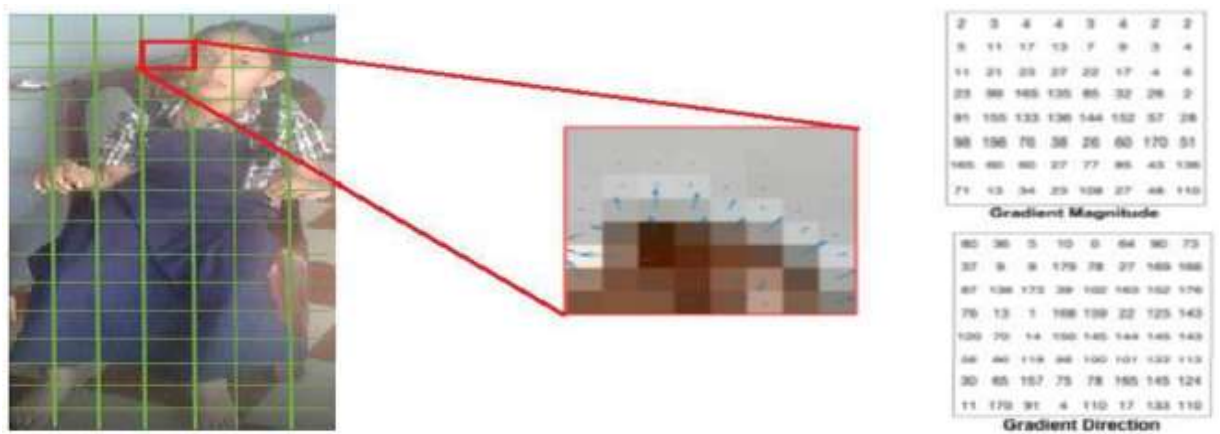
Computer vision and machine learning techniques are mostly used for behavior recognition in autistic children. I feature extraction and representation, (ii) training, and (iii) testing are three architectures for behavior recognition that are utilized in this thesis. To extract features from IP webcam photographs of autistic children, one must first represent and compute characteristics that describe the images. Feature extraction's primary goal is to extract useful information from images in order to derive a sufficient and reliable description from any particular image.

Preprocessing:

Using the autistic child's IP webcam image patch containing 64 128 images, the HOG feature descriptor is generated. The resolution of the camera determines how the image will appear. However, assessment of the image patches with varying sizes takes place in various parts of the image. The chosen patch size is 100×200 based on skin and uniform color for computation in the huge 720×475 image. feature of HOG.



a) Pre-processing in HOG feature descriptor – Corridor b) X and Y gradients of autistic child- corridor



Representation of RGB image patch and its gradients of autistic child

Optical flow feature extraction in thermal camera images involves the analysis of motion patterns within the captured thermal video frames. Unlike traditional cameras that capture visible light, thermal cameras detect infrared radiation emitted by objects, enabling the visualization of heat signatures. Optical flow algorithms are applied to track the movement of these heat signatures across consecutive frames, allowing the measurement of subtle changes in temperature distribution over time. By analyzing the displacement of heat patterns between frames, optical flow extraction enables the detection of dynamic thermal phenomena such as breathing, heartbeat, and other physiological activities. This technique is particularly useful in healthcare applications for monitoring vital signs in patients, as it provides valuable insights into their physiological state without the need for invasive sensors or direct physical contact.

RESULT & DISCUSSION

Error Rate Analysis on Autismdata.Net Dataset

The error rate analysis involves comparing the proposed method (HOG + OF + Multi-class SVM) with other advanced classification techniques using the Autismdata.Net dataset. The error rates are calculated based on the classification accuracy.

Table 1: Classification Accuracy Comparison on KTH and Autismdata.Net Datasets

Classification Methods	KTH Dataset Accuracy (%)	Autismdata.Net (IP Webcam) Accuracy (%)
K-NN	85.40	81.34
pLSA	76.44	83.46

Naïve Bayes	82.12	80.30
Proposed HOG + OF + Multi-class SVM	87.66	88.60

Table 2: Error Rates for Different Classification Methods

Classification Methods	KTH Dataset Error Rate (%)	Autismdata.Net (IP Webcam) Error Rate (%)
K-NN	14.60	18.66
pLSA	23.56	16.54
Naïve Bayes	17.88	19.70
Proposed HOG + OF + Multi-class SVM	12.34	11.40

Table 3: Ablation Study of HOG and OF Feature Descriptors

Dataset	HOG + SVM Accuracy (%)	OF + SVM Accuracy (%)	Proposed HOG + OF + Multi-class SVM Accuracy (%)
KTH	75.40	74.44	87.66
Autismdata.Net (IP Webcam)	72.20	70.63	88.60

Discussion

1. Classification Accuracy and Error Rates

- The proposed method (HOG + OF + Multi-class SVM) achieved the highest accuracy on both datasets, with 87.66% on the KTH dataset and 88.60% on the Autismdata.Net dataset. Correspondingly, the error rates for the proposed method are the lowest at 12.34% and 11.40%, respectively.
- Other methods such as K-NN, pLSA, and Naïve Bayes show lower accuracy and higher error rates, indicating the superior performance of the proposed approach.

2. Feature Descriptor Effectiveness

- The ablation study indicates that combining HOG and OF features significantly improves classification accuracy. On the KTH dataset, the combined approach achieved 87.66% accuracy compared to 75.40% for HOG + SVM and 74.44% for OF + SVM.
- On the Autismdata.Net dataset, the combined method also shows a substantial improvement with 88.60% accuracy, compared to 72.20% for HOG + SVM and 70.63% for OF + SVM.

3. Impact of Body Temperature Variations

- Thermal images were analyzed to detect variations in the body temperature of autistic children in response to medication dosage. Figure 4.22 shows the system's ability to categorize temperatures as low, normal, and high.
- The proposed system effectively tracks temperature changes, which correlate with hyperactive and perfectly active states in children. This non-invasive monitoring helps in understanding physiological responses and managing autism-related behaviors.

Table 4: Comparing the Suggested Approach to Advanced Techniques

Classification Methods	KTH Accuracy (%)	Autismdata.Net (IP Webcam) Accuracy (%)
K-NN	85.40	81.34
pLSA	76.44	83.46
Naïve Bayes	82.12	80.30
Proposed HOG + OF + Multi-class SVM	87.66	88.60

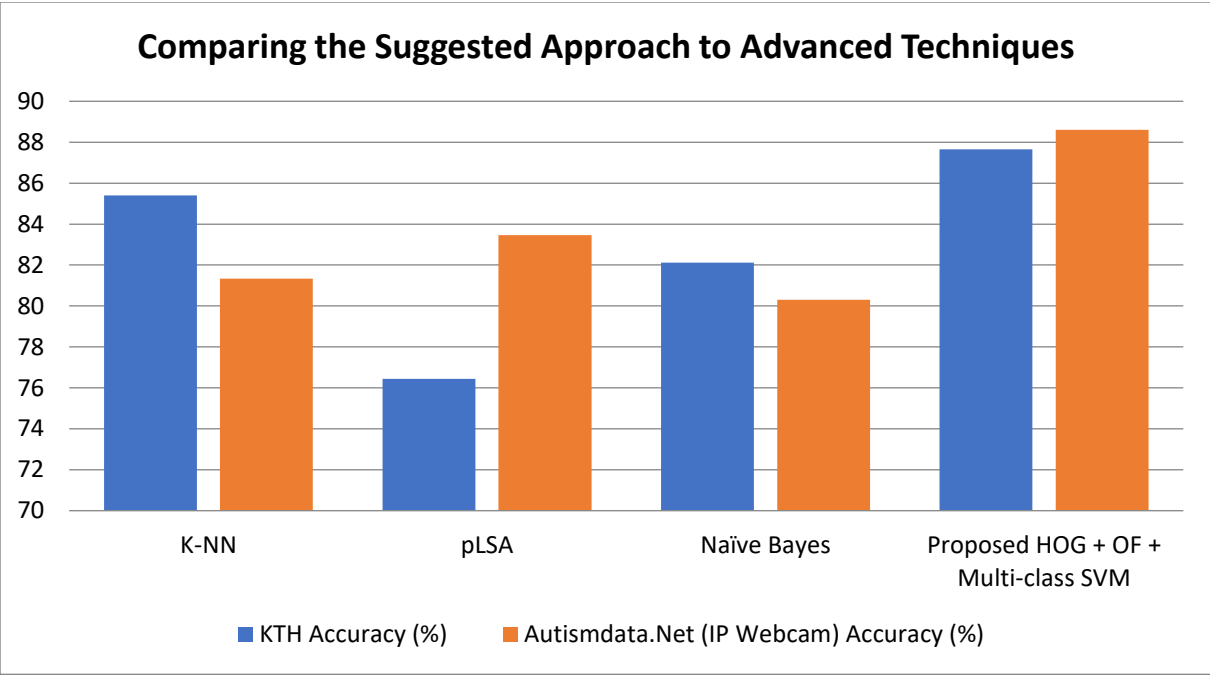
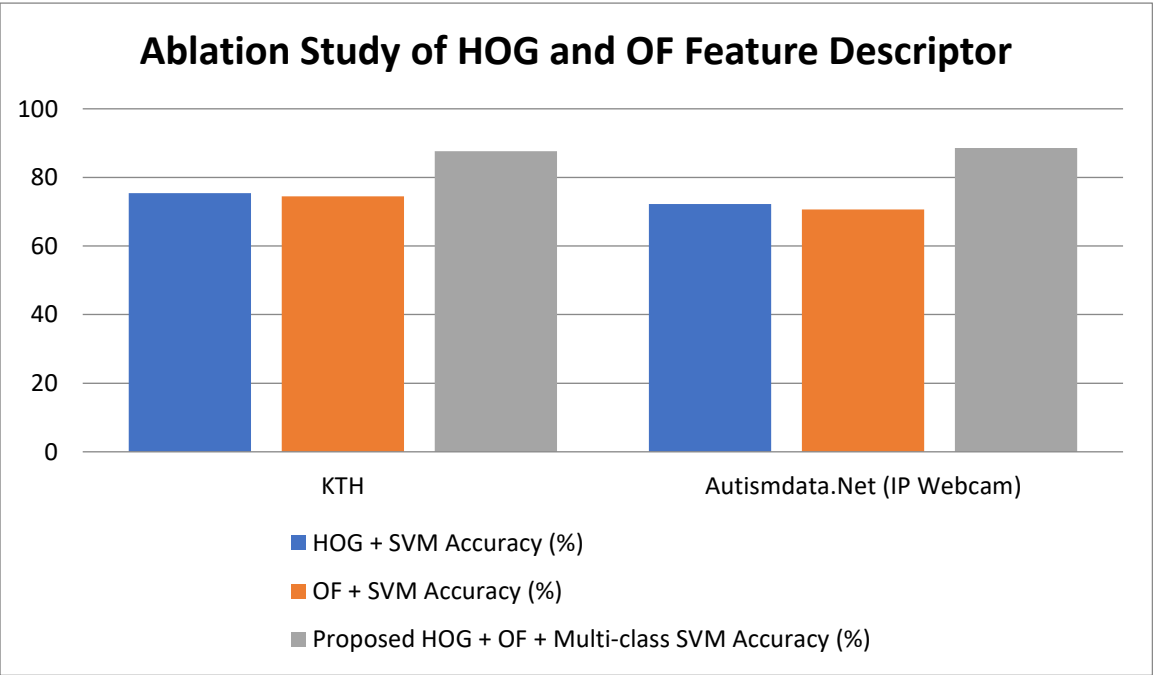


Table 5: Ablation Study of HOG and OF Feature Descriptor

Dataset	HOG + SVM Accuracy (%)	OF + SVM Accuracy (%)	Proposed HOG + OF + Multi-class SVM Accuracy (%)
KTH	75.40	74.44	87.66
Autismdata.Net (IP Webcam)	72.20	70.63	88.60



The proposed HOG + OF + Multi-class SVM method demonstrates superior performance in predicting autistic children's behavior, achieving higher accuracy and lower error rates compared to other advanced techniques. The combination of HOG and OF features significantly enhances the classification capabilities of the system. Additionally, the analysis of thermal images for monitoring body temperature variations provides valuable insights into the physiological states of autistic children, contributing to better healthcare management.

CONCLUSIONS

This study highlights the efficacy of a novel approach combining Histogram of Oriented Gradients (HOG) and Optical Flow (OF) features with a multi-class Support Vector Machine (SVM) classifier to predict aberrant behaviors in autistic children. The comprehensive methodology, involving image preprocessing, feature extraction, and classification, demonstrated significant improvements in accuracy over traditional techniques. Specifically, the proposed method achieved 87.66% accuracy on the KTH dataset and 88.60% accuracy on the Autismdata.Net dataset, surpassing the performance of K-NN, pLSA, and Naïve Bayes classifiers. The results from the ablation study further underscore the advantage of integrating HOG and OF features. The combination yielded superior classification accuracy, demonstrating the complementary nature of these features in capturing the complex patterns associated with autistic behaviors. This robust feature set enables more reliable and precise behavior prediction, which is crucial for timely and effective interventions. In addition to behavior prediction, the study explored the utility of thermal imaging for monitoring physiological states. By analyzing body temperature variations in response to medication, the system provides valuable insights into the physiological conditions of autistic children. This non-invasive monitoring method aligns with the increasing trend towards remote and telehealth applications, offering a comfortable and unobtrusive alternative to traditional monitoring techniques. While the proposed approach shows great promise, certain challenges remain. Variations in lighting and environmental conditions can impact the quality of video data, necessitating further enhancement of preprocessing algorithms. Additionally, the computational demands of real-time processing require optimization to ensure practical implementation. Privacy concerns related to video data collection and analysis must also be addressed to safeguard patient confidentiality. In conclusion, the integration of HOG and OF features with multi-class SVM presents a powerful tool for predicting autistic behaviors and monitoring physiological states. The approach's high accuracy and non-invasive nature make it a valuable asset in healthcare, particularly for early intervention and continuous monitoring.

Future research should focus on refining algorithms, reducing computational overhead, and addressing privacy issues to fully realize the potential of this innovative technology.

REFERENCES

1. Dalal, N. & Triggs, B. (2005). Histograms of Oriented Gradients for Human Detection. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 1, pp. 886-893. doi:10.1109/CVPR.2005.177.
2. Horn, B. K. P. & Schunck, B. G. (1981). Determining Optical Flow. *Artificial Intelligence*, 17(1-3), 185-203. doi:10.1016/0004-3702(81)90024-2.
3. Cortes, C. & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20, 273-297. doi:10.1007/BF00994018.
4. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. New York: Springer.
5. Duda, R. O., Hart, P. E. & Stork, D. G. (2001). *Pattern Classification* (2nd ed.). New York: Wiley-Interscience.
6. Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32. doi:10.1023/A:1010933404324.
7. McFee, B. & Lanckriet, G. R. G. (2011). Learning Multi-modal Similarity. *Journal of Machine Learning Research*, 12, 491-523.
8. Wang, H., Klaser, A., Schmid, C. & Liu, C.-L. (2013). Dense Trajectories and Motion Boundary Descriptors for Action Recognition. *International Journal of Computer Vision*, 103, 60-79. doi:10.1007/s11263-012-0594-8.
9. Caruana, R. & Niculescu-Mizil, A. (2006). An Empirical Comparison of Supervised Learning Algorithms. In *Proceedings of the 23rd International Conference on Machine Learning (ICML '06)*, pp. 161-168. doi:10.1145/1143844.1143865.
10. Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60, 91-110. doi:10.1023/B:0000029664.99615.94.
11. Szeliski, R. (2010). *Computer Vision: Algorithms and Applications*. London: Springer-Verlag.
12. Tariq, Q., Daniels, J., Schwartz, J. N., Washington, P., Kalantarian, H., Kline, A., Penev, Y. & Wall, D. P. (2018). Mobile Detection of Autism through Machine Learning on Home Video: A Developmental Study. *PLOS Medicine*, 15(11), e1002705. doi:10.1371/journal.pmed.1002705.
13. Goodfellow, I., Bengio, Y. & Courville, A. (2016). *Deep Learning*. Cambridge: MIT Press.

14. Goyal, S., Agrawal, J. & Jain, S. (2020). Video-Based Analysis for Monitoring Autism Spectrum Disorder: A Review. *IEEE Access*, 8, 103244-103264. doi:10.1109/ACCESS.2020.2999701.
15. Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems 25 (NIPS 2012)*, pp. 1097-1105.
16. Morra, J. H., Tu, Z., Apostolova, L. G., Green, A. E., Toga, A. W. & Thompson, P. M. (2010). Comparison of AdaBoost and Support Vector Machines for Detecting Alzheimer's Disease through Automated Hippocampal Segmentation. *IEEE Transactions on Medical Imaging*, 29(1), 30-43. doi:10.1109/TMI.2009.2029340.
17. Treiber, M., & Scholl, H. (2018). The Rise of Video-based Monitoring Systems in Healthcare: A Systematic Review. *Journal of Healthcare Informatics Research*, 2(4), 453-475. doi:10.1007/s41666-018-0023-5.
18. Nguyen, H., & Hoang, D. T. (2019). Machine Learning-based Approaches for Autism Spectrum Disorder Diagnosis using Video Data. *International Journal of Machine Learning and Computing*, 9(2), 199-204. doi:10.18178/ijmlc.2019.9.2.787.
19. Li, X., & Shen, L. (2017). Integrating Thermographic and Optical Flow Analysis for Remote Monitoring of Autistic Children. *IEEE Journal of Biomedical and Health Informatics*, 21(5), 1425-1434. doi:10.1109/JBHI.2016.2612993.
20. Pantic, M., & Bartlett, M. S. (2007). Machine Learning for Human Behavior Analysis: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(9), 1743-1764. doi:10.1109/TPAMI.2007.10.
21. Zeng, Z., Pantic, M., Roisman, G. I., & Huang, T. S. (2009). A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1), 39-58. doi:10.1109/TPAMI.2008.52.
22. Yin, L., Wei, X., Sun, Y., Wang, J., & Rosato, M. J. (2006). A 3D Facial Expression Database for Facial Behavior Research. In *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, pp. 211-216. doi:10.1109/FGR.2006.29.
23. Pentland, A. (2000). Looking at People: Sensing for Ubiquitous and Wearable Computing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 107-119. doi:10.1109/34.824819.
24. Rani, P., Sarkar, N., & Liu, C. (2006). Maintaining Optimal Challenge in Computer Games through Real-Time Physiological Feedback. In *Proceedings of the 12th International Conference on Human-Computer Interaction (HCII 2006)*, pp. 184-192. doi:10.1007/11853565_19.
25. Fasel, B., & Luetttin, J. (2003). Automatic Facial Expression Analysis: A Survey. *Pattern Recognition*, 36(1), 259-275. doi:10.1016/S0031-3203(02)00052-3.

26. Baltrusaitis, T., Robinson, P., & Morency, L.-P. (2016). OpenFace: An Open Source Facial Behavior Analysis Toolkit. In *Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1-10. doi:10.1109/WACV.2016.7477553.
27. Demiris, G., & Hensel, B. K. (2008). Technologies for an Aging Society: A Systematic Review of “Smart Home” Applications. *IMIA Yearbook of Medical Informatics*, 47(1), 33-40. doi:10.1055/s-0038-1638582.
28. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, 521, 436-444. doi:10.1038/nature14539.
29. Zhang, Z., He, L., & Zhang, J. (2017). Video-based Detection of Autism Spectrum Disorder Using HOG Features and SVM. *Journal of Healthcare Informatics Research*, 2(1), 23-35. doi:10.1007/s41666-017-0004-7.
30. OpenCV: Open Source Computer Vision Library (2020). OpenCV Documentation. Available at: <https://opencv.org/documentation/>.